Supplementary Information for "The civilizing process in London's Old Bailey"

Sara Klingenstein [*], Tim Hitchcock [†] and Simon DeDeo [*] [‡] [§]

[*]Santa Fe Institute, Santa Fe, NM 87501, USA, [†]Department of History, University of Sussex, Brighton, BN1 9RH, United Kingdom, [‡]School of Informatics and Computing, Indiana University, 901 E 10th Street, Bloomington, IN 47408, USA, and [§]To whom correspondence should be addressed. Email: sdedeo@indiana.edu

Summary of Supplementary Information

This supplementary information file presents detailed characterization of the Old Bailey over the one-hundred and fifty four years of the data we use.

We first test for two major potential systematic influences: (1) those related to changes in the Old Bailey's jurisdiction, and (2) those related to overall trial volume and number of words.

We then present an account of the ratio between indictments for violent and non-violent crimes and the overall volume of trial accounts, connecting it to both formal changes in Old Bailey jurisdiction and shifts in cultural attitudes.

We then characterize the changing nature of indictments at the Old Bailey, and connect this to both cultural and legal practice.

An additional supplementary analysis is then presented reflecting a result discussed in the main article, *viz.*, that the civilizing process seen at the large-scale level can also be seen at the indictment level. We describe and present detailed graphical and statistical analyses of these sub-processes.

We then present a sample trial from our dataset, with the spoken word testimony highlighted.

Finally, we compare the compare the British English corpus from Google N-grams with the word frequencies found in the Old Bailey Corpus, for the most relevant synonym sets by Partial KL. This allows a year-by-year comparison of word usage as a check against the possibility of systematic and unaccounted-for drift in the use of language at large.

S1: Jurisdiction Tests

Here we examine a crucial "deflationary" explanation for the long-timescale shifts in JSD: that administrative changes in the nature of crimes allocated to the Old Bailey could mimic the genre distinction that this work interprets as being due to large-scale shifts in cultural and bureaucratic norms. If, for example, a trial type is simply shifted to a different court, this could lead to an apparent creation or disappearance of a trial genre without a concomitantly significant change in actual cultural or legal practice. The major administrative shifts that could lead to such an effect in the data are shown as vertical lines in Fig. 1: 1834, when the jurisdiction was extended to the whole of England, and 1855, when less-serious crimes were reallocated to Magistrates courts.[1]

We find that these two dates do not leave a significant, distinguishable signature on the time series, testing this observation via a simple model-selection argument. We consider two models for the post-1778 Jensen-Shannon divergence: the first with two discrete steps at the 1834 and 1855, and a second, linear secular-increase model that does not require dividing the data into epochs. The log-likelihood of the data given the model allows us to reject the administrative-shift model, for the secular-increase model. The log of the maximum-likelihood ratio, $\ln\left[P_{\max}(\text{linear}|\text{data})/P_{\max}(\text{stepwise}|\text{data})\right]$ is equal to +99 for the category level and +102 for the synonym set level. In both cases, the evidence in favor of the secular-

increase model is at a significance, or p value, much less than 10^{-12}.

This result holds even if we use a generalized three-epoch model, where the two dates of transition are allowed to float freely, independently at each coarse-graining level, and without regard to known administrative or historical events. The secular increase model is still strongly preferred ($\Delta\mathcal{L} = +49$ for the category level and $\Delta\mathcal{L} = +29$ at the synonym set level; *i.e.*, still preferred at $p < 10^{-12}$), despite the fact that this generalized model now has five free parameters compared to the two-parameter linear process.

S2: Trial Statistics

We consider only trials with semantically identifiable spoken-word testimony. While most trials contain spoken testimony, cases settled by plea bargain, and hence resulting in a bare record of charge, guilt and punishment, do not. As a result, trials of this sort, which become relatively commonplace after approximately 1820, form the bulk of the data dropped from the OBC, leaving a total of 112,485 trials, and a total of 2.3×10^7 semantically significant words. Figure S3 shows the number of trails available to us once these cuts are made, as well as the average number of words per trial. Not all words are classified in Roget's; determiners such as "the", for example, are without direct lexical semantics. Our analysis works solely with the words assigned semantic significance.

In estimating the relevant information-theoretic quantities of the trials from the data, it is necessary to bias-correct the naively-measured values, and to determine their error ranges; we do this by means of the statistical bootstrap, which preserves the coarse-graining relationship at better than 10^{-4} for the levels of over-sampling we achieve in our data [3]. We do this bootstrap correction by resampling on a trial-by-trial level, a more conservative method that takes into account systematic variation due to within-trial correlations.

The most serious possible contaminant of the strong result we find in Figure 1 is the possibility that over time, as trials became longer and more detailed, more distinctive language was introduced. Fig. S3 allows us to check that the shifts seen in Fig. 1 are not simply due to changes in trial length.

While trial length is positively correlated with Jensen-Shannon divergence (Pearson cross-correlation of 0.52 ± 0.09), the relationship between JSD and trial length is clearly non-monotonic. For example, the Old Bailey of the 1780s devotes more words per trial than in the 1820s, but it is this latter period that shows a strong signal of the genre distinction. It is likely that the same bureaucratic and cultural processes that drive the emergence of the genre distinction are also affected by the time and resources devoted to individual trials, and thus that correlations in the quantities are to be expected.

---

[1] "The Criminal Justice Act", 18 & 19 Vic. c.126 (1855) established summary jurisdiction on a clearly defined basis, allowing people charged with minor theft and other offences to be convicted by two justices. This act was amended only slightly by the Summary Jurisdiction Act, 42 & 43 Vic. c.49 (1879). For a narrative account of the evolution of the bureaucracy of criminal justice in these decades see Ref. [1]. For a statistical approach to the impact of this legislation see Ref. [2].

However, a simple model that assumes a monotonic relationship between trial length and distinctiveness can not explain the results of Fig. 1.

S3: Indictments: Violent vs. Non-Violent

This section considers the change in the number, and ratio, of violent and non-violent cases, and discusses possible influences on our main results.

From 1760 to 1913, the Old Bailey's trials are a snapshot of how the criminal behavior of London was managed by the bureaucracy. Fig. S4 shows the rise, from 1800 to 1840, in the overall number of non-violent indictments (as found in the OBC), and its subsequent decline after 1840. Because trials in the Old Bailey are mostly for non-violent crimes, Fig. S4 has roughly the same shape as the dotted line, reflecting the number of trials overall, in Fig. S3.

Fig. S5 shows the same counts for violent crimes, and Fig. S6 the ratio of violent to non-violent crimes. Between 1760 and 1834, the overall ratio of violent to non-violent crimes declines very slightly—from a high of around 10% in the 1760s to around 5% by 1834 (statistical significance for the negative slope is formally at $+21\sigma$ given bootstrap errors). A major part of this decline results from the fall in the per capita rate of violent crime, which traces the arc of the civilizing process in real-world behavior.

While many violent crimes certainly went unreported, and some jurisdictional and legal changes affected the precise mix of crimes that came to the Old Bailey, the most serious crimes, and in particular, murder, would have always come to the Old Bailey in this period. The murder rate has been recognized as a standard metric for changing levels of violence. While pre-1801 population figures for London are subject to some error, between 1760 and 1815 the population of London nearly doubled, from roughly three quarters of a million to 1.4 million in 1815, rising to three million in 1861, and to over seven million in 1911. The largely unchanging number of murder cases (Fig. S7) heard at the Old Bailey each year translates into a substantial per capita decline. This statistic is complicated by the contemporaneous increase in prosecutions for the lesser crime of manslaughter. This could result from either new perceptions of culpability in accidental death, or a growing leniency on the part of the prosecution, in cases that would previously have been tried as murder. But even in the case of manslaughter, the number of prosecutions per capita of the population was effectively falling. For readers interested in detailed discussion of the overall level of crime in London, see Refs. [4, 5, 6].

In addition to the increasing safety of London, as evidenced by the decline in the per capita murder rate, another source of the decline in the ratio between violent and non-violent crime arises from an increase in prosecutions for non-violent offences. The growth in this category of indictments is more rapid than even the most liberal estimates of population growth in London between 1760 and 1834, and cannot be attributed to this growth alone. In part, the increase could be due to a rise in actual crime, but is more likely to have resulted from the removal of capital punishment for most forms of theft between 1834 and 1842. This transition could potentially have encouraged victims to proceed with formal prosecutions, rather than relying on extra-judicial punishments and negotiations. An effect of this sort can be seen earlier in the data when, in 1808, Parliament removed the death penalty for pickpocketing, leading to a concomitant increase in prosecutions.[2] Meanwhile, the overall number of trials grew substantially, reflecting a criminal justice system responding to a rapidly growing population.

In 1834, the Old Bailey was given jurisdiction over the whole of England for serious crimes only; this leads to a systematic increase in the ratio of violent to non-violent crimes that can not be attributed to changes in behavior or legal culture. The fraction of violent crimes undergoes a sudden rise, driven in part by the increasing numbers of serious crimes committed outside of London that ended up in the Old Bailey, but more significantly by the decline in the number of prosecutions. The post-1840s decline in the number of non-violent crimes is cemented by the jurisdictional shift of the Criminal Justice Act in 1855 that gave greater scope to the Magistrates' courts. We examine the characteristics of this latter effect in the following section.

With these provisos in mind, the violent to non-violent trial ratio can be seen to divide in to three distinct phases: (1) slightly declining, from 1760 to 1834; (2) a transition in 1834 to about 1860 driven by jurisdictional change; (3) a second reasonably stable phase. The main underlying social processes that drove these dynamics include, for example, not just an overall decline in the murder rate, but also increasingly limited use of capital punishment; both of which are classic components of the civilizing process. Meanwhile, our statistical tests for a three-epoch model, described in the main text, allow us to reject a simple model where the JSD is driven by systematic effects associated with the indictment ratio (Fig. S6), or by the 1840s transition in the overall number of trials for violent indictments (Fig. S5).

S4: Indictments: Internal Structure

This section describes the changing composition of the Old Bailey indictments in greater detail, including relevant legal-historical details and cultural processes.

Our data contains a total of sixty distinct indictment types; the prevalence of which varies widely, both within a particular year, and over the course of the period reflected in the data. Because of the changing volume of trials over the course of the data, it is more revealing to consider the average yearly indictment fractions, rather than the overall fractions of indictments. Fig. S8 shows this distribution on a log-linear scale to give a sense of how much an "average" year at the Old Bailey concentrates on different types of crime.

A small number of indictments account for the majority of cases; the top eleven indictments account for 90% of all non-violent offences (the top five, in the case of violent indictments). Concern for property dominates both the violent and non-violent indictments; trials for violent theft ("highway robbery" and "robbery") average 54% of a year's docket for violent offences, while trials for theft (ranging from larceny to pocketpicking, poaching, and extortion) make up 84% of the non-violent docket. Other forms of violence—wounding, murder, manslaughter, infanticide and rape are the next five most common—are subdominant concerns; meanwhile, forms of deception not involving direct theft (counterfeiting, forgery, fraud, bigamy and perjury) are the next most common indictment types for non-violent offences, with other non-violent crimes of concern to the state (including non-violent sexual offences, return from transportation, libel, sedition, and illegal abortion).

The diversity of indictments—here quantified as the entropy of the distribution over indictments—is shown in Fig. S9. In contrast to the class of violent indictments, which maintains a reasonably constant diversity, the non-violent indictments are increasingly diverse over time. The transition

---

[2] 48 Geo. III, c. 129, s. 2 (1808); see Ref. [7].

is well-marked by the two major jurisdictional shifts; it begins around 1834, a period of definitional change, and finishes around 1855. Again, our statistical tests for a three-epoch model, allow us to reject a simple model where the JSD is driven by this two-phase diversity structure.

Fig. S10 shows the changing make-up of the list of non-violent indictments. The most visually striking shift, in 1827, is purely procedural in origin. Indictments for grand and petty larceny were combined by act of Parliament into "simple larceny".[3] One can also see the effect of the 1808 pickpocketing statute mentioned above. The upper red band corresponding to the pickpocketing indictments undergoes rapid expansion correlating with the implementation of new statute law. A similar pattern can be seen in the indictment class including "stealing from master", following an act in 1823 (upper blue band).

Fig. S11 shows the changing make-up of the list of violent indictments. The two most striking shifts are (1) the near-complete disappearance of the "highway robbery" violent theft category after 1829, and (2) the increasing representation of "wounding" as an indictment.

In a similar fashion to the consolidation of larceny indictments into a broader class, phenomenon (1) appears to stem from a purely formal change in legal categorization; this happened in two stages, with the first shift happening in 1827 (under the same acts consolidating larceny offences) and then in 1828;[4] these two acts together eliminated a range of previous indictment subcategories and established new definitions—a set of practices that took a year or so to propagate to the judges in the courtroom itself and leading to the recategorization of highway robbery under other forms of violent theft (here collected into the Old Bailey's violent theft/robbery category).

Phenomenon (2) appears as a much longer-term shift towards prosecutions for non-fatal assault. This appears to be a signature, at the crime-categorization level, of the same civilizing process that we detect at the fine-grained level of recorded speech. Meanwhile, despite the fact that the violent crime docket begins to include crimes of lesser severity (wounding as opposed to killing or violent theft), the overall distinguishability of violent and non-violent indictments continues to increase.

S5: The Civilizing Process by Indictment

Based on findings in the previous two sections, this section examines the civilizing process at the more fine-grained level of indictment classes. We find strong evidence for the same civilizing process found in the system as a whole, further confirming and elaborating the results of the main text.

In particular, we study the increasing trial distinctiveness (Jensen-Shannon Distance) between groups of indictments within each of the violent and non-violent classes. Because some indictments are very rare (see Fig. S8), we collect indictments into six major subgroups; three subgroups for violent indictments, and three for the non-violent indictments.

The non-violent subgroups are **Theft**, **Deception**, and **Other**.

In the OBO categorization system, **Theft** contains grandLarceny, pettyLarceny, simpleLarceny, theftFromPlace, burglary, pocketpicking, stealingFromMaster, receiving, animalTheft, housebreaking, shoplifting, embezzlement, mail, extortion, gameLawOffence, breakingIntoPlace and the "other" theft category; **Deception** contains fraud, forgery, perjury, bankruptcy, the "other" deception category and bigamy and coiningOffences; the **Other** category collects a number of offences that do not clearly fall under either prop-

erty crime or simple deception, and includes arson, libel, returnFromTransportation, pervertingJustice, conspiracy, taxOffences, sodomy, keepingABrothel, concealingABirth, seditiousLibel, treason, seducingAllegiance, seditiousWords, religiousOffences, illegalAbortion, piracy, habitualCriminal, vagabond and "miscellaneous" other.

The violent subgroups are **Violent Theft**, (non-fatal) **Assault**, and **Killing**.

In the OBO categorization system, **Violent Theft** contains robbery, highwayRobbery, and the "other" violentTheft category. **Assault** contains wounding, assault, threateningBehaviour, rape, kidnapping, riot, indecentAssault and assaultWithSodomiticalIntent. **Killing** contains murder, manslaughter, infanticide, killing/pettyTreason, and the "other" killing category.

We can then ask (for example) whether **Violent Theft** and **Theft** show the same increasing distinctiveness found for the violent and non-violent categories as a whole. Table S1 shows the nine possible pairwise comparisons for both the coarse-grained and fine-grained semantic levels, and we show graphical results in Figs. S12 through S14.

Both **Violent Theft** and **Assault** show significant evidence ($> 5\sigma$) for a linear increase in trial distinguishability against all three non-violent indictment categories from 1778 to 1913. The most significant signal is the increasing distinctiveness of violent and non-violent theft, and the largest rate of increase is found for the distinction between assault and deception and violent theft and deception. Among all the pairs, violent theft and theft are the hardest to distinguish in Jensen-Shannon distance—as is to be expected, given that the entire distinction hinges on the involvement of violence—and this distinction undergoes significant growth over the course of the data.

Category (V)	Category (NV)	Slope (bits/decade)	std dev
Violent Theft	**Theft**	0.0011	+20.1
	Deception	0.0022	+14.6
	Other	0.0012	+8.5
Assault	**Theft**	6×10^{-4}	+6.2
	Deception	0.0026	+14.2
	Other	6×10^{-4}	+3.6
Killing	Theft	-3×10^{-4}	−3.1
	Deception	0.0013	+6.8
	Other	-1×10^{-4}	−0.8

Fig. S1. Linear model parameters and significance (standard deviations away from zero) for indictment class pairs at the "Category Level" coarse-graining (equivalent to top of Fig. 1 of the main text). Indictment group pairs with significant evidence of increasing distinctiveness ($> 5\sigma$) are marked in **bold**.

Category (NV)	Category (NV)	Slope (bits/decade)	std dev
Theft	**Deception**	-8×10^{-4}	−8.1
Theft	**Other**	-6×10^{-4}	−5.8
Deception	Other	$+6 \times 10^{-4}$	3.8

Fig. S2. Linear model parameters and significance (standard deviations away from zero) for indictment class pairs, at the "Category Level" coarse-graining, within the non-violent indictment class.

[3] 7 & 8 Geo. IV, c. 29: An Act for Consolidating and Amending the Laws of England Relative to Larceny and other Offences Connected Therewith

[4] 9 Geo. IV, c. 31: An Act for consolidating and Amending the Laws of England Relative to Offences Against the Person

Killing, by contrast, shows evidence for this process only in the case of deception. Comparison with Figs. S12 through S14 suggests that part of the reason is that strong distinctions *already* exist in 1778—it is only in the 1900s, for example, that trials for violent theft and trials for non-violent theft become as distinct as trials for theft and trials for killing. This suggests that what we are seeing in the years available to us here is the further "perfection" of a longer-running process.

This pattern is in contrast to the development of distinctions within the non-violent indictment class; as shown in Table S2 and Fig. S15, non-violent indictment classes themselves become less distinguishable over time.

S6: The Trial of Mary Hall

The Trial of Mary Hall for Grand Larceny, 28 October 1801.

MARY HALL was indicted for feloniously stealing, on the 19th of October, five guineas, the property of John Martin.

JOHN MARTIN sworn.—I am a labourer: On the 19th of October, I was going up Back-lane, Rosemary-lane, about two o'clock in the day; I met with a friend, and got a little too much; I was returning home, and met with the prisoner in Wellclose-square; I went home with her; she asked me what I would give her for a bed that night.

Q. What time was it?—A. I cannot exactly tell; I agreed to give her two shillings; we slept together till between five and six in the morning; I had my breeches under my head, and in them a purse, containing five guineas; when I waked in the morning, between five and six, I searched, and found my money safe; I got up to evacuate, and while I was doing that, she got the money from me; I saw her take the money out of my breeches; I then laid hold of her, and had a struggle with her, to get the money back again, and she cried out, and a man came into the room with a knife in his hand, and swore, if I did not stand back, he would run the knife into my heart; I still kept in contention with her, and then a woman, that he called his wife, came up with a pair of tongs, and hit me with the tongs on my back; I took the tongs from her, and then the man said, if I did not lay down the tongs, he would surely run the knife through my heart; I was afraid of being murdered, and I put on my breeches, my waistcoat, and my jacket, and there came a black man behind me, he struck me upon my back, and I made resistance till I got out into the street; then the people gathered about, and they threw my shoes and stockings out at window, and I put them on; I then went to Lambeth-street, and got an officer.

RICHARD OSMAN sworn.—I am an officer of Lambeth-street, Whitechapel: On Tuesday, the 20th of October, the prosecutor came to me, and said, he was robbed of five guineas; that he had been used very ill, and asked me to go with him to apprehend the man that had robbed him; he said, he saw the girl with the money in her hand; I went, and apprehended her; I searched her, and found upon her one shilling, and some halfpence; she said, she knew nothing about it, she had no other money; he told me, there was an old man drew a knife upon him.

Q. Who is that man?—A. It turned out to be a man of the name of Stinson; he saw him in the House of Correction; he told me exactly the same story before I apprehended the girl, that he has now told your Lordship.

Q. I suppose he was quite sober at that time?—A. Yes. Q. How came Stinson not to be committed?—A. Before the girl had her last examination, he was taken in execution, and sent to the House of Correction.

Prisoner's defence. The man came into the room to my assistance, with his breakfast in his hand.

GUILTY, aged 29.

Second Middlesex Jury, before Mr. Common Serjeant.

Total length: 563 words. Total spoken: 524 words. Total lexical semantics: 147 words. Old Bailey Online Trial ID t18011028-44; last accessed 23 November 2013.

Our classification of violent indictments, following OBO definitions: offenceCategory={kill or violentTheft} or offenceSubcategory={assault [assault, assaultWithIntent, assaultWithSodomiticalIntent and indecentAssault], riot, threateningBehaviour, rape, kidnapping, or wounding}; all other offence categories (including, for example, the grand larceny indictment above) are classed as "non-violent" indictments. A very small fraction of trials (181 out of 112,485) have multiple indictments for both violent and non-violent offenses; we classify these trials as "violent."

S7: Old Bailey Corpus vs. Google N-Grams

The Old Bailey Corpus is a particularly long-baseline sample of spoken English, produced in the particular context of the criminal justice bureaucracy. It is worth comparing the synonym set categorization of this data against a larger corpus, both to see large-scale trends in the English language as well as to check for the possibility of biases that might exist due to the changing relevance of Roget's categorization. The top twenty-four most relevant synonym sets for our analysis are shown at the end of this section. There we compare the synonym-set frequency in the OBC to that found for the top ten words of those sets in the Google N-Grams corpus (British English, Ref. [8]). Since the Google N-Grams corpus normalizes by all words (as opposed to by semantically significant words), frequencies are roughly a factor of five lower; we scale the Google corpus by this factor to make visual comparison easier.

As can be read from those figures, most of the variation in the Old Bailey Corpus is only weakly correlated to variation in the use of British English in the Google corpus as a whole (Pearson cross-correlation, 0.10 ± 0.02). Further, in most cases, the synonym set density in the Google corpus shows significantly less relative variance than in the OBC (11% *vs.* 29%, $\pm 1\%$), suggesting that the bureaucratic system itself is undergoing greater change than the culture as a whole.
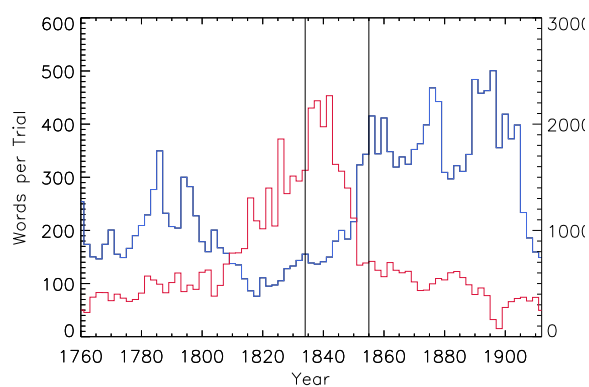
Fig. S3. Semantically significant words per trial (blue line) and total number of trials per year (red line) in the OBC.
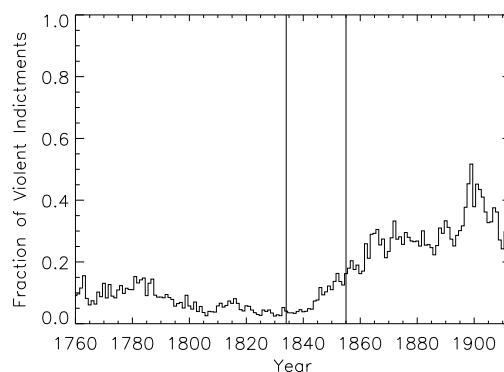

Fig. S6. Overall ratio of trial indictments, violent to non-violent, in the data.
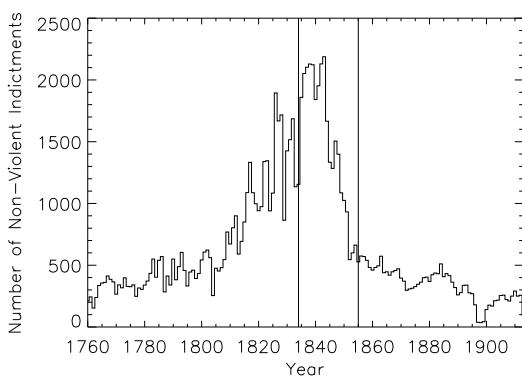

Fig. S4. Total number of indictments for non-violent crimes in the corpus.
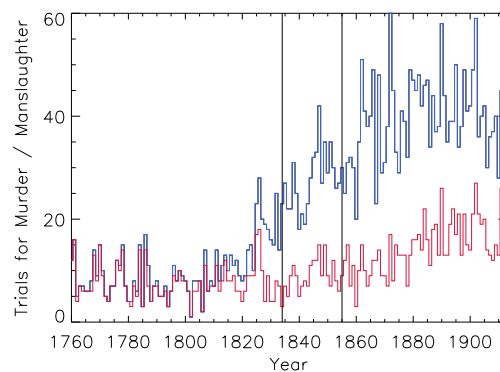

Fig. S7. Total number of trials in the OBO (not OBC) for murder and manslaughter (blue line) and murder alone (red alone).
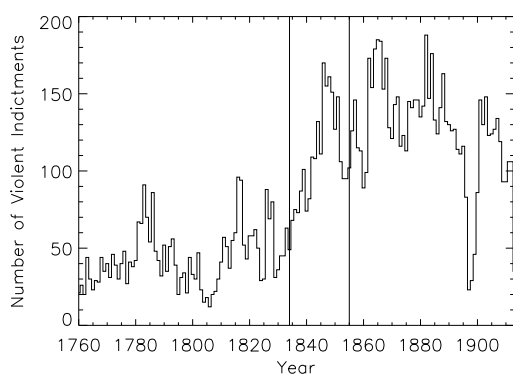

Fig. S5. Total number of indictments for violent crimes in the corpus.
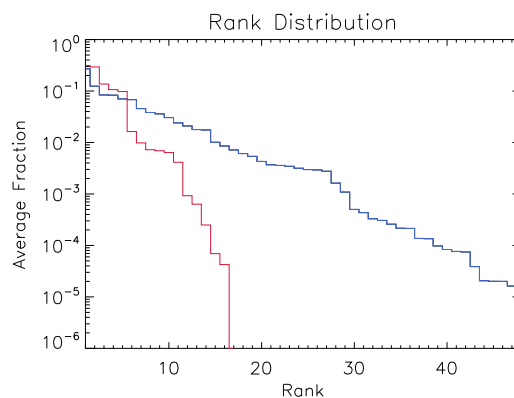
Fig. S8. Distribution of offence category-subcategory pairs within the OBC; broken out by indictments for non-violent offences (48 indictments total; blue line) and violent offences (16 total; red line).
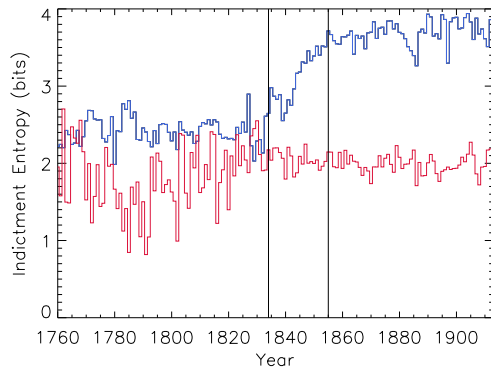
Fig. S9. Diversity of indictment types, broken out into violent (red line) and non-violent (blue line) categories. While the entropy for the violent indictment class remains roughly constant over 150 years, the non-violent indictments become increasingly diverse over time, undergoing a distinct phase shift between 1830 and 1850.

Fig. S11. Composition of the violent indictment class. From bottom to top, violent theft/robbery (red), violent theft/highway robbery (green), wounding (blue), murder (yellow), manslaughter (purple), simple assault (brown); the remainder in white.

Fig. S10. Composition of the non-violent indictment class. From bottom to top, grand larceny (red), petty larceny (green), simple larceny (blue), theft from place (yellow), coining offences (purple), burglary (brown), pocketpicking (red), fraud (green), "stealing from master" (blue), forgery (yellow), receiving stolen goods (purple); remainder in white. Simple larceny combined the grand and petty larceny cases by statue in 1827. In addition to the 1834 and 1855 points, we overlay the 1808 date, when pocketpicking was removed from the list of capital offences.

Fig. S12. The Jensen-Shannon Divergence for violent theft and (1) theft (blue line), (2) deception (red line), and (3) all other non-violent offences (green line). In all three cases, strong evidence exists for the same civilizing process observed in the overall evolution of the Old Bailey. As expected, the hardest to distinguish pair is violent theft and theft (the JSD is lower than the other two pairs), but still shows significant ($p \ll 10^{-5}$) evidence for increasing divergence over time.
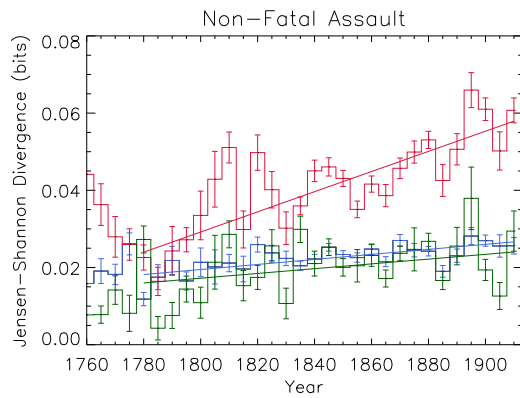
Fig. S13. The Jensen-Shannon Divergence for non-fatal assault and (1) theft (blue line), (2) deception (red line), and (3) all other non-violent offences (green line). Strong evidence exists for an increasing divergence over time between both assault and theft, and assault and deception.
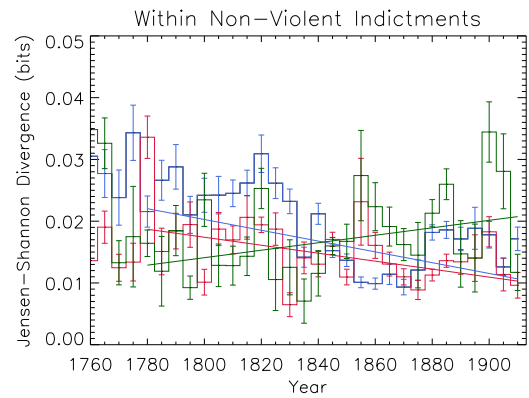
Fig. S15. The Jensen-Shannon Divergence for indictment types within the non-violent class: (1) theft–deception (blue line), (2) theft–other (red line), and (3) deception–other (green line). In contrast to the violent–non-violent genre emergence, deception and "other" become increasingly *less* distinct from the larger theft class over time, and only marginal ($< 4\sigma$) evidence exists for an increasing distinction between deception and other.
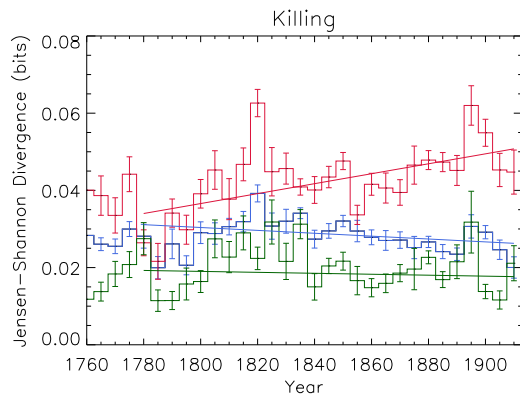
Fig. S14. The Jensen-Shannon Divergence for offences involving killing (murder, manslaughter, infanticide, treasonous killing) and (1) theft (blue line), (2) deception (red line), and (3) all other non-violent offences (green line). Here, only deception and killing show strong evidence ($> 5\sigma$) for increasing divergence over time.
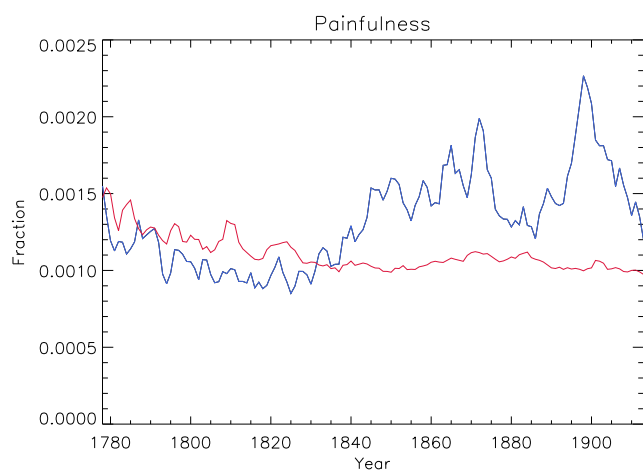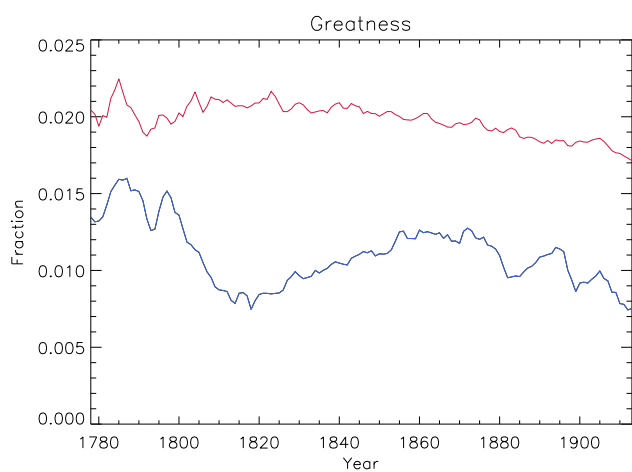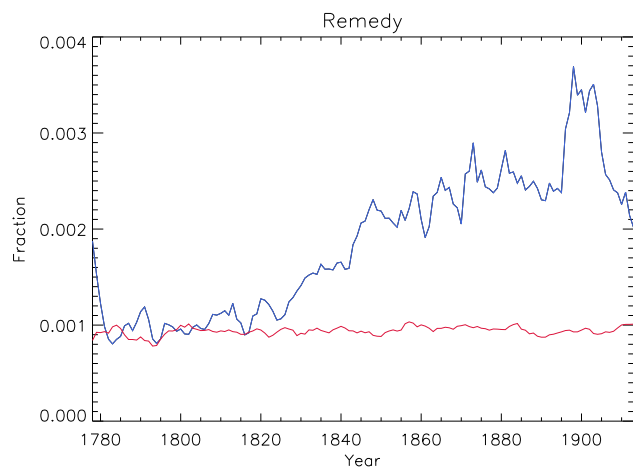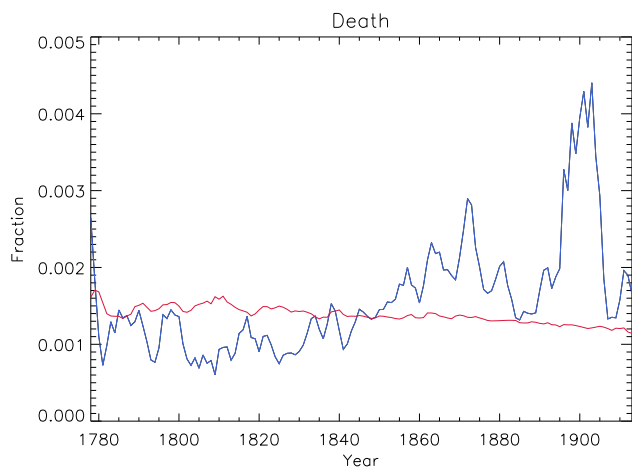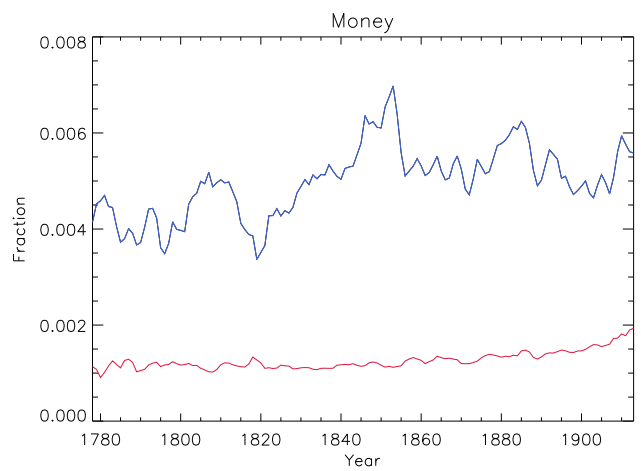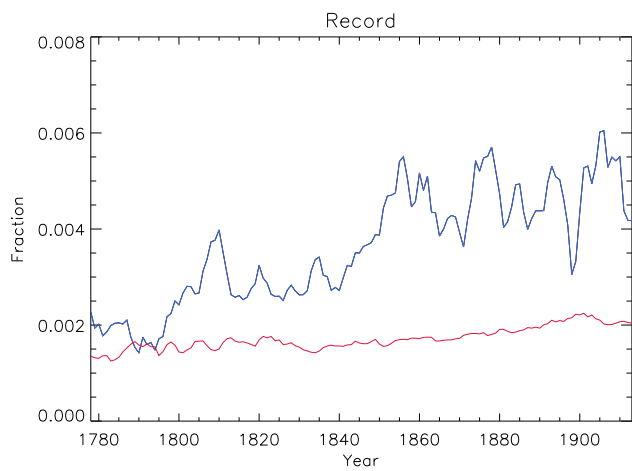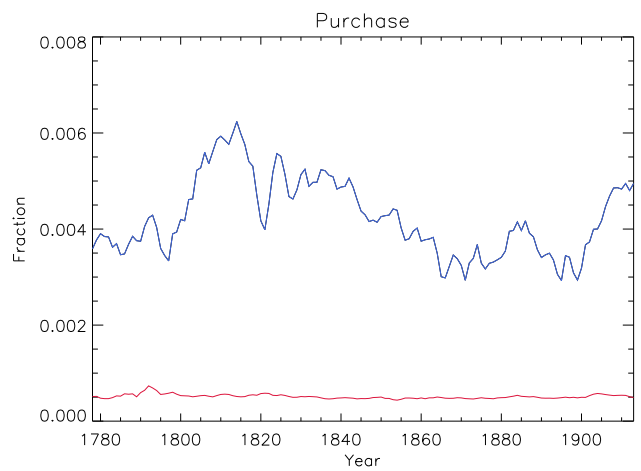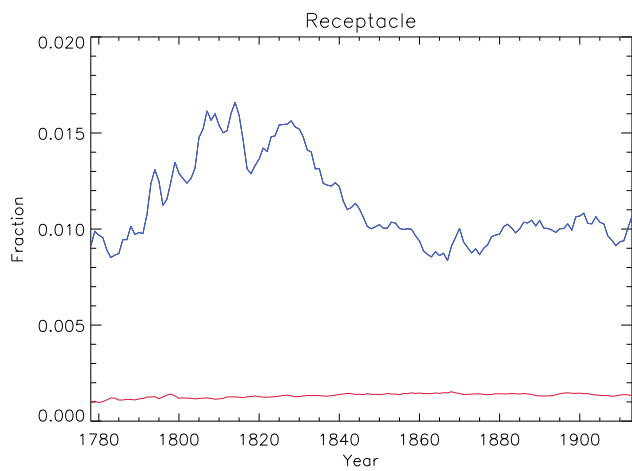
Fig. S16. Overall abundance of words. Old Bailey Corpus (blue line) *vs.* Google N-Grams British English database (red line, $\times 5$) for the top synonym sets associated with violent offences, Fig. 3(a) of the main text.
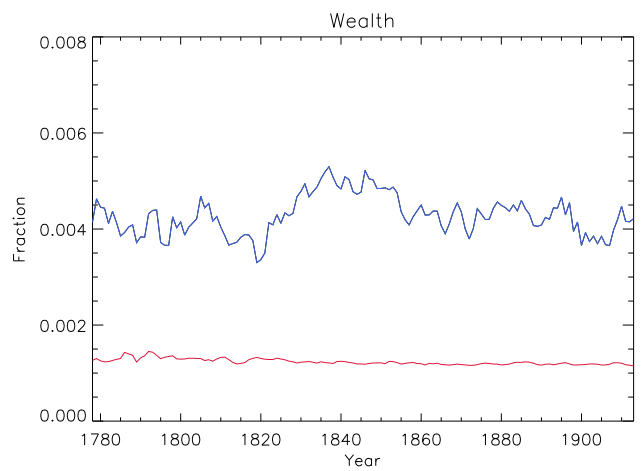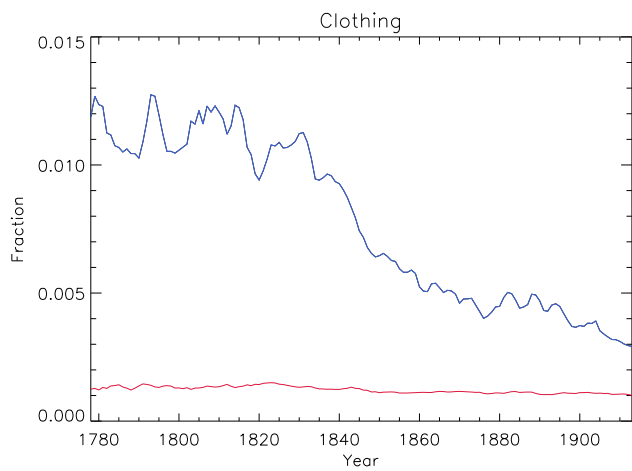
Fig. S17. (continued).
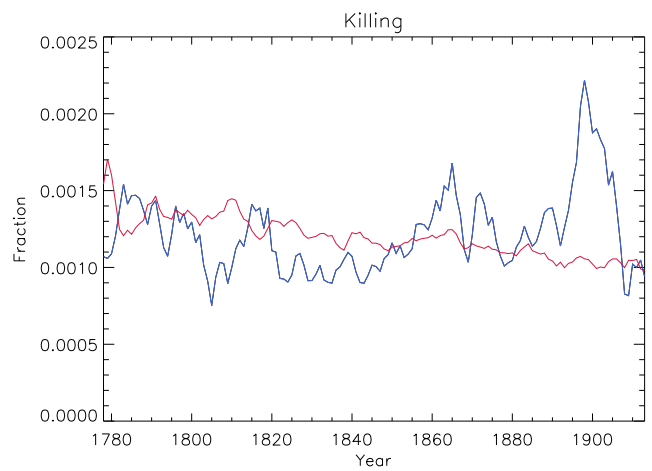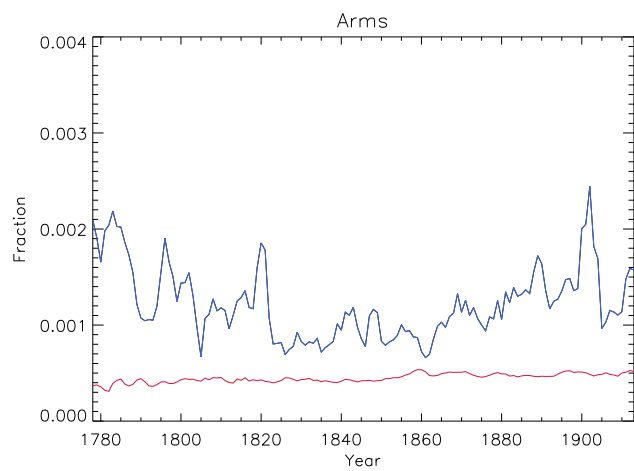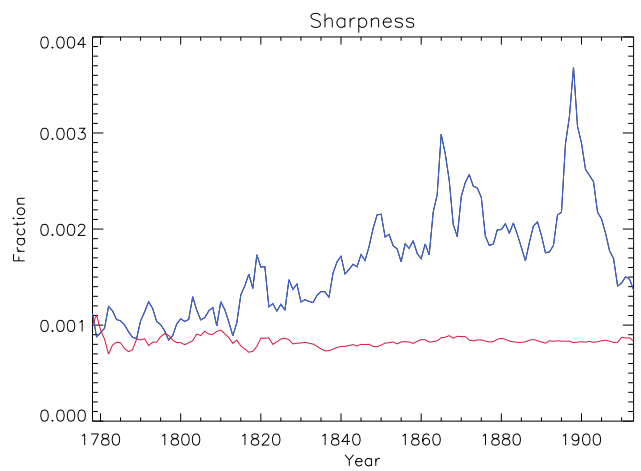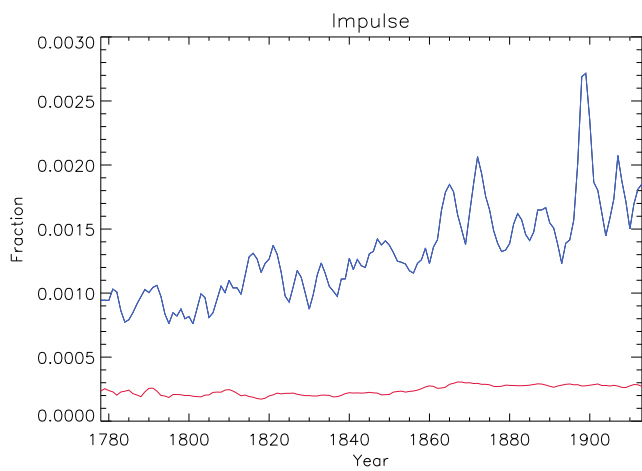
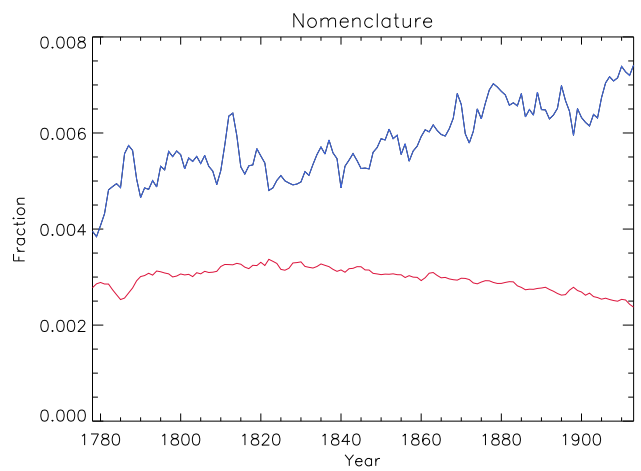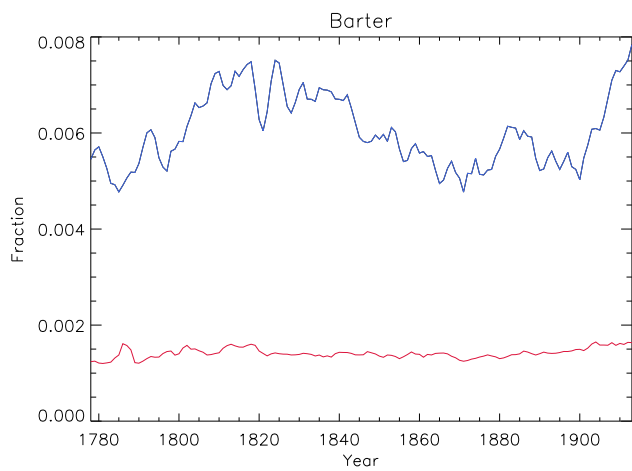**Fig. S18.** (continued).

Fig. S19. (continued).
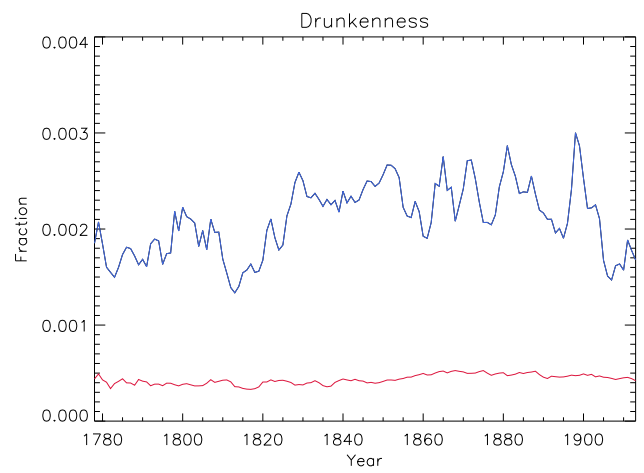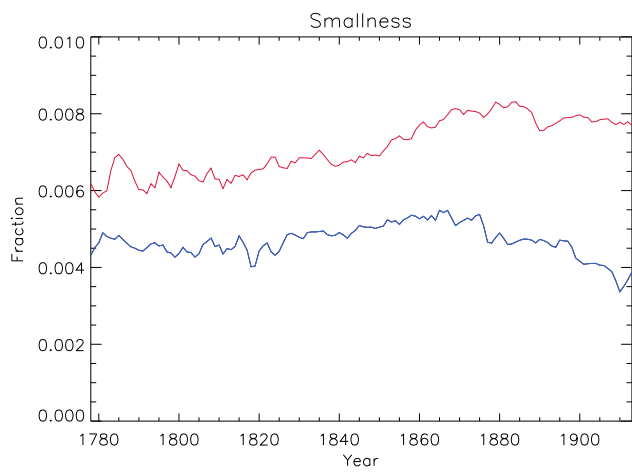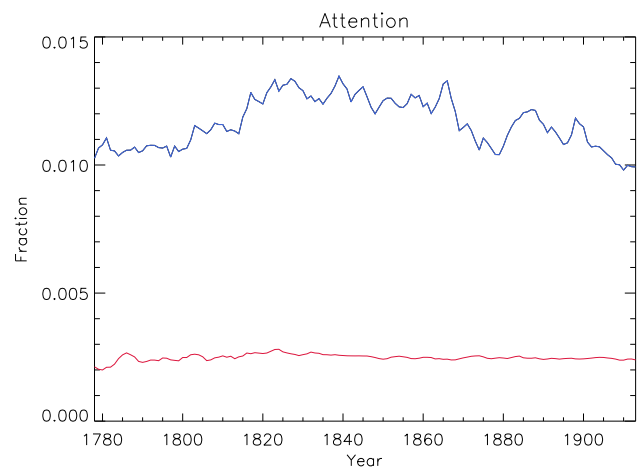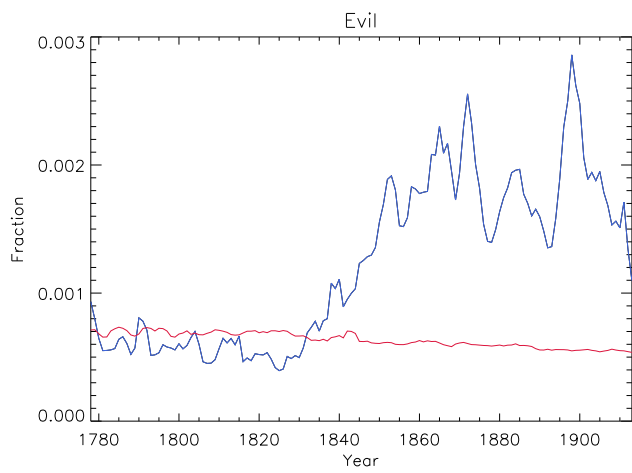
Fig. S20. (continued).

Fig. S21. (continued).

Fig. S22. Overall abundance of words. Old Bailey Corpus (blue line) *vs.* Google N-Grams British English database (red line, ×5) for the top six synonym sets associated with non-violent offences, Fig. 3(b) of the main text.

Fig. S23. (continued).

Barter

Nomenclature

Clothing

Wealth

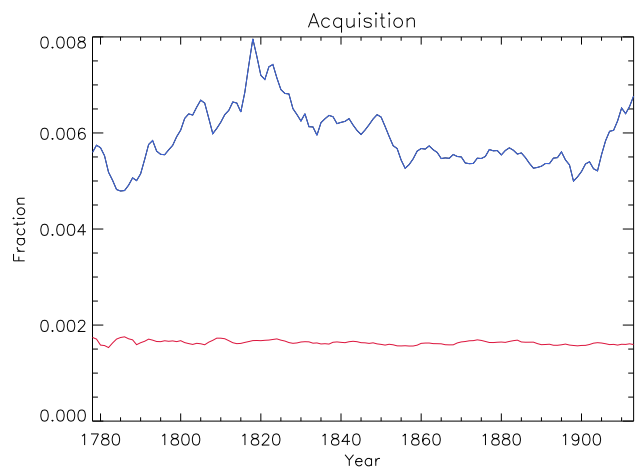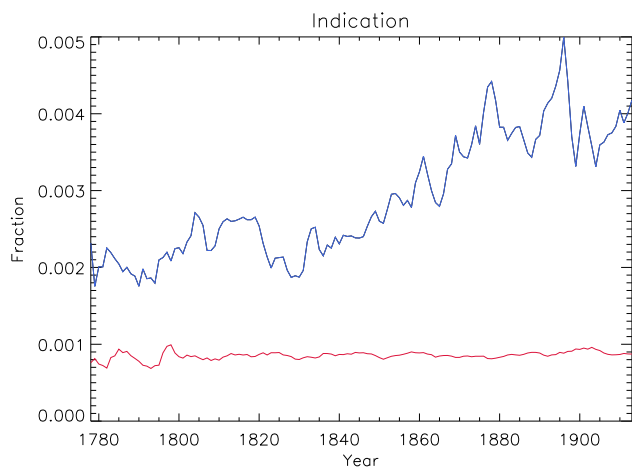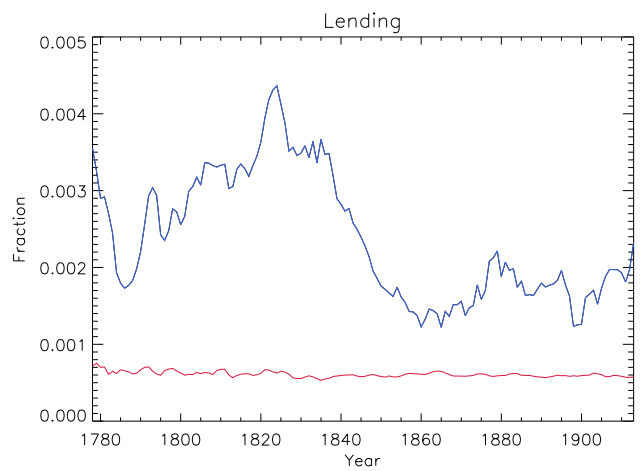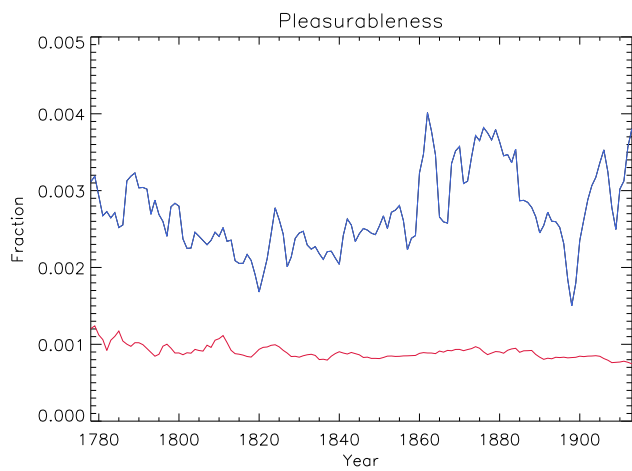**Fig. S24.** (continued).

Fig. S25. (continued).

Fig. S26. (continued).

Fig. S27. (continued).

1. Emsley C (2010) Crime and society in England, 1750–1900 (Longman, London) Fourth edition, Ch. 8.

2. Williams CA (2000) Counting crimes or counting people: some implications of mid-nineteenth century british police returns. Crime, Histoire & Sociétés 4:77–93.

3. DeDeo S, Hawkins RX, Klingenstein S, Hitchcock T (2013) Bootstrap methods for the empirical study of decision-making and information flows in social systems. Entropy 15:2246–2276.

4. Eisner M (2003) Long-term historical trends in violent crime. Crime and Justice 30:83–142.

5. Cockburn JS (1991) Patterns of Violence in English Society: Homicide in Kent 1560-1985. Past & Present 130:70–106.

6. Gurr TR (1981) Historical Trends in Violent Crime: A Critical Review of the Evidence. Crime and Justice 3:295–353.

7. Beattie J (2012) The First English Detectives: The Bow Street Runners and the Policing of London, 1750-1840 (OUP Oxford) Pg. 211.

8. Michel JB, et al. (2011) Quantitative Analysis of Culture using Millions of Digitized Books. Science 331:176–182 British English Corpus; actual search at http://bit.ly/1e5AvqY; last accessed 11 December 2013.